

Metabolic isotopic labeling of (deoxy)ribose enables dual neutral loss scanning for profiling nucleic acid modifications

Shao-Qin Rong^{1,2,3,4,5,†}, Zhi-Han Jin^{1,5,†}, Dong-Rui Yin^{1,5,†}, Jun Yao¹, Guo-Liang Xu^{1,2}, Ping Zhu^{6,7,*}, Hai Gao^{5,*}, Dan Zhou^{8,*}

¹Shanghai Key Laboratory of Medical Epigenetics, Institutes of Biomedical Sciences, Medical College of Fudan University, Shanghai 200032, China

²CAS Key Laboratory of Epigenetic Regulation and Intervention, Shanghai Key Laboratory of Molecular Andrology, Shanghai Institute of Biochemistry and Cell Biology, Center for Excellence in Molecular Cell Science, Chinese Academy of Sciences, Shanghai 200031, China

³Advanced Molecular Pathology Institute of Soochow University and SANO, Suzhou 215123, China

⁴Suzhou SANO Precision Medicine Ltd, SANO Medical Laboratories, Suzhou 215123, China

⁵Shanghai Xuhui Central Hospital, Zhongshan-Xuhui Hospital, Fudan University, Shanghai 200237, China

⁶Guangdong Cardiovascular Institute, Guangdong Provincial People's Hospital (Guangdong Academy of Medical Sciences), Southern Medical University, Guangzhou, Guangdong 510100, China

⁷Guangdong Provincial Key Laboratory of Pathogenesis, Targeted Prevention and Treatment of Heart Disease, Guangzhou Key Laboratory of Cardiac Pathogenesis and Prevention, Guangzhou, Guangdong 510100, China

⁸Center for Medical Research and Innovation, Shanghai Pudong Hospital, Institutes of Biomedical Sciences, Medical College of Fudan University, Shanghai 201399, China

*To whom correspondence should be addressed. Email: danzhou@fudan.edu.cn

Correspondence may also be addressed to Hai Gao. Email: gaohai@fudan.edu.cn

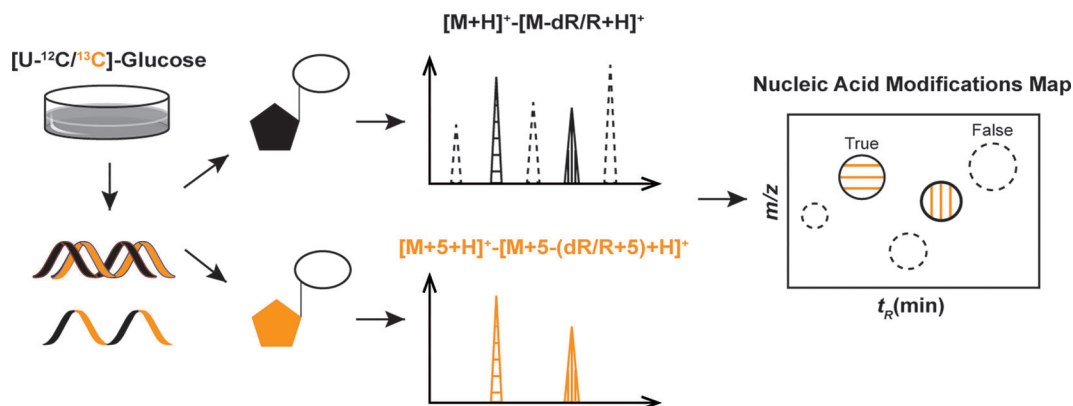
Correspondence may also be addressed to Ping Zhu. Email: tanganqier@163.com

[†]The first three authors should be regarded as Joint First Authors.

Abstract

Modifications on DNA and RNA, even at trace levels, play critical roles in diverse biological processes, yet their accurate quantification and discovery of novel modified constituents remain challenging. Here, we present a metabolic (deoxy)ribose-labeling approach integrated with multiple reaction monitoring mass spectrometry (MRM-MS) to enable sensitive and untargeted detection of nucleic acid modifications. In this approach, [U-¹³C]glucose is used to generate an ~1:1 ratio of ¹³C₅-labeled and unlabeled deoxyribose moieties *in vivo*, followed by MRM-MS acquisition of neutral-loss transitions corresponding to both isotopic forms. This isotopic pairing facilitates the differentiation of endogenous nucleosides from contaminants and allows confident assignment of authentic nucleoside signals. Applying this method to mouse embryonic stem cells, we detected rare nucleoside species such as 5-formylcytosine and 5-carboxycytosine, present at frequencies as low as one in 10⁶–10⁷ bases. In contrast, peaks assigned to *N*⁶-methyladenine (6mA) lacked a labeled counterpart, suggesting that previously reported 6mA in mammalian DNA may arise from RNA misincorporation or artifacts introduced during the processing of isolated DNA. Analysis of formaldehyde-treated DNA revealed several previously unreported adducts, including *N*⁴-hydroxymethyl-5-hydroxymethylcytosine (4hm5hmC, or dihmC). Collectively, this (deoxy)ribose-labeling strategy provides a robust and sensitive platform for untargeted nucleoside profiling and the discovery of uncharacterized nucleic acid modifications.

Graphical abstract



Received: February 11, 2025. Revised: February 6, 2026. Accepted: February 6, 2026

© The Author(s) 2026. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License

(<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.

Introduction

DNA and RNA modifications play crucial roles in regulating cell identity and activity by modulating gene expression without altering the underlying genetic sequences. To date, over 170 distinct modifications have been identified in RNA [1], and >17 have been reported in DNA [2]. Among vertebrate genomes, 5-methylcytosine (5mC) installed by DNA methyltransferases (DNMTs) is the most abundant DNA modification. The oxidized derivatives of 5mC—5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC), and 5-carboxycytosine (5caC)—are generated by the ten-eleven translocation (TET) family of dioxygenases and occur at levels that are two to four orders of magnitude lower than that of the 5mC precursor. Additionally, rare modifications such as 5-hydroxymethyluracil (5hmU) and 5-formyluracil (5fU), likely produced by TET enzymes, are present at sub-ppm levels in mammalian genomes [3, 4]. Despite their low abundance, these 5mC oxidation products appear to be indispensable for gene regulation, such as suppressing aberrant splicing (e.g. 5hmC) [5] and modulating RNA polymerase activity (via 5fC and 5caC) [6, 7]. Nevertheless, the full spectrum of modified nucleic acids and their profound roles in fine-tuning transcriptional and epigenetic processes remain largely unexplored, leaving open the possibility that undiscovered rare modifications could further advance our understanding of gene regulation and cellular function.

Beyond the nucleoside modifications well-characterized in vertebrate genomes, a diverse array of distinct nucleobase modifications has been identified across other domains of life. In prokaryotes and lower eukaryotes, N^6 -methyladenine (6mA) and N^4 -methylcytosine (4mC) are prevalent [2, 8], alongside hypermodified bases like 5-glyceryl-methylcytosine (5gly-mC) [9] and β -d-glucopyranosyloxymethyluracil (base J) [10]. Even greater chemical diversity has been observed in bacteriophages, where nucleobases conjugated to amino acids, polyamines, and sugars have been detected [11]. These modifications play epigenetic roles or function in host-pathogen defense and counter-defense mechanisms [2, 11]. The discovery of such a wide spectrum of nucleobase modifications has spurred efforts to explore novel DNA modifications in higher eukaryotes, particularly in mammals and in specific organelles such as mitochondria and chloroplasts. However, the existence of certain modifications—such as 5mC reported in mitochondrial DNA and 6mA, 4mC, and 4acC in the mammalian genome—remains contentious, largely due to technical limitations and potential artifacts. Although some studies have reported their detection at trace levels using high-throughput sequencing and mass spectrometry [12–18], others using the same methods have failed to confirm their presence [19–22]. These inconsistencies highlight the complexity of detecting rare modifications and underscore the need for more robust strategies to distinguish genuine DNA modifications from potential artifacts. Identifying new and biologically relevant nucleobase modifications remains a critical frontier in life sciences, with profound implications for understanding gene regulation, cellular processes, and disease etiology.

The identification and quantification of nucleobase modifications are critically influenced by the choice of analytical platform. While traditional techniques such as thin-layer chromatography have enabled the discovery of novel base derivatives, they require radioactive labeling (^{32}P or ^3H) and offer limited resolution and structural information [23–25].

In contrast, liquid chromatography-mass spectrometry (LC-MS/MS) has emerged as a powerful and versatile tool for both the detection and structural elucidation of DNA modifications [9, 24–27]. Multiple reaction monitoring (MRM), a scanning mode unique to QqQ-MS, enables sensitive detection of predefined precursor-product ion transitions. By configuring the first quadrupole (Q1) and the second quadrupole (Q3) to a constant 116-Da mass offset—corresponding to the characteristic neutral loss of a deoxyribose moiety during fragmentation—MRM can be adapted to scan for transitions from $[\text{M} + \text{H}]^+$ to $[\text{M} + \text{H} - 116]^+$ across a selected m/z range [28]. This approach, later termed pseudo-neutral loss (PNL) or stepped MRM [28, 29], enables untargeted detection of deoxynucleosides with broad coverage and operational efficiency. However, the unit resolution of QqQ instruments (typically 0.5–1 Da) limits the specificity of such scans. A 116-Da neutral loss is not exclusive to deoxyribose and may arise from structurally unrelated compounds, raising the possibility of false positives. This inherent limitation in mass selectivity presents a major challenge for accurate annotation of nucleoside-derived signals. To fully harness the screening efficiency and reduce the risk of misidentification in MRM analyses, it is imperative to implement orthogonal strategies that enhance molecular specificity and thereby enable the reliable identification of genuine nucleoside-derived features.

To address the challenge of distinguishing true nucleoside-derived ions in QqQ-MS analysis, we developed a dual neutral loss scanning strategy based on *in vivo* labeling of nucleosides with a 1:1 ratio of ^{13}C - and ^{12}C -isotopes in the (deoxy)ribose. By culturing mouse embryonic stem cells (mES) with $[\text{U-}^{13}\text{C}]$ glucose, a subset of (deoxy)ribose units in DNA and RNA molecules were heavy-isotope labeled, establishing a near-equimolar distribution compared with the light isotopologues. MRM monitors neutral-loss transitions from both ^{13}C - and ^{12}C -incorporated forms, and peaks exhibiting matched intensities between the isotopologues are interpreted as authentic nucleoside species. This isotopic pairing significantly enhances molecular specificity while maintaining high sensitivity, allowing reliable detection of rare modifications present at frequencies as low as one per 10^6 – 10^7 bases, exemplified by the ready detection of 5fC and 5caC in mES DNA. It also enabled discrimination between endogenous modifications and artifactual species, as shown by the reinterpretation of previously reported 6mA as a likely contaminant. Application of this method to formaldehyde (FA)-treated genomic DNA revealed several novel nucleoside species undetected by conventional MS workflows, including N^4 -hydroxymethyl-5-hydroxymethyl-cytosine (hereafter referred to as dihmC), a labile DNA modification with an estimated half-life of ~25 min. Together, our findings establish the use of both ^{13}C - and ^{12}C -incorporated (deoxy)ribose as an effective orthogonal strategy to improve the authenticity and specificity of nucleoside identification, offering a robust platform for high-throughput discovery of previously unrecognized DNA modifications.

Materials and methods

Reagents

The following chemicals were used for LC-MS/MS: methanol, acetonitrile (ACN), formic acid, acetic acid (AA), ammonium

formate, ammonium acetate, and ammonium bicarbonate, all of which were of LiChropur™ grade and purchased from Sigma–Aldrich (St. Louis, MO, USA). Malic acid (MA), 3,5-di-tert-butyl-4-hydroxytoluene (BHT), and deferoxamine mesylate salt (DFOA) were also obtained from Sigma–Aldrich. The deaminase inhibitor tetrahydrouridine (THU) and erythro-9-(2-hydroxy-3-nonyl)adenine hydrochloride (EHNA) was sourced from MedChemExpress (Princeton, NJ, USA). All deoxynucleotides and [U-¹³C]glucose were procured from ZZBio (Shanghai, China).

Cell culture and isotope labeling

The wild-type (WT) and DNMT triple knockout (DNMT TKO) mES cells were maintained on 0.1% gelatin-coated plates and cultured in DMEM without glucose (GIBCO), supplemented with 20 mM [U-¹³C]glucose and 5 mM [U-¹²C]glucose (Sigma), 15% FBS (GIBCO), 2 mM L-glutamine, 0.1 mM nonessential amino acids (GIBCO), 1 mM sodium pyruvate (GIBCO), 0.1 mM β-mercaptoethanol (β-Me), 1 × Pen/Strep, 3 μM CHIR99021, 1 μM PD0325901, and 1000 U/ml mouse LIF (Sigma–Aldrich). Cells were plated at a density of 1 × 10⁶ cells per 10 cm dish and cultured under these conditions for 48 h.

DNA and RNA purification and enzymatic hydrolysis

Genomic DNA was extracted following a modified protocol based on the method reported by the Thomas Carell group [30], using the QIAGEN AllPrep DNA/RNA Mini Kit supplemented with protectants. This kit allows completion of the isolation process within 1 h, thereby mitigating the degradation issue. The isolation and hydrolysis of genomic DNA and RNA often require the use of various chemical and enzymatic reagents, which can inadvertently damage nucleobases or degrade labile modifications [30–32]. Antioxidants, including BHT (200 μM), DFOA (200 μM), β-Me (0.1%), along with the deaminase inhibitors THU (200 μM) and EHNA (200 μM), were added into all buffers to reduce background oxidation or deamination (Supplementary Fig. S1A). Elution was performed using ddH₂O supplemented with 200 μM BHT and DFOA, and the eluates were stored at –80°C for long-term preservation.

To minimize incubation time for complete hydrolysis of DNA or RNA, the digestion protocol was as follows. A total of 100 μg genomic DNA or RNA was dissolved in 100 μl ddH₂O. To this, 2 μl of 50 × ND1 buffer (1 M NH₄Ac, pH 5.3 and 50 mM MgCl₂) was added, followed by 100 U of Nuclease P1 (New England BioLabs) and 10 U of DNase I (Takara) or 100 μg RNaseA. The reaction mixture was incubated at 37°C for 30 min to break down these oligonucleotides into mononucleotides (Supplementary Fig. S1B and C). Subsequently, 5 μl of 10 × CIAP buffer (Takara) and 30 U of Calf Intestinal Alkaline Phosphatase (CIAP, Takara) were added to the mixture, followed by incubation at 37 °C for 1 h to ensure ≥99.9% dephosphorylation of mononucleotides (Supplementary Fig. S1D and E). Before LC-MS/MS analysis, the reaction mixture was filtered using a 3K ultrafiltration tube (PALL) to remove macromolecular substances such as proteins, and the filtrate was collected for analysis.

Chromatographic conditions

The LC-MS/MS analysis was conducted using a coupled system comprising a 1290 Infinity II UHPLC (Agilent) and a 6500 + QTRAP mass spectrometer (AB SCIEX).

To enable the detection of a wide range of nucleosides with good separation and high sensitivity, a reversed-phase liquid chromatography (RPLC) and a hydrophilic interaction liquid chromatography (HILIC) were employed. The Waters ACQUITY Premier HSS T3 column (RPLC, 1.8 μm, 2.1 × 150 mm) was operated at a flow rate of 0.3 ml/min with a column temperature of 40°C, and the ACQUITY Premier BEH Amide column (HILIC, 1.7 μm, 2.1 × 150 mm) was operated at a flow rate of 0.4 ml/min with a column temperature of 45°C. A selection of 16 deoxyribonucleosides and 14 ribonucleosides, both modified and unmodified, was used for assessing chromatographic conditions. Previous studies have shown that ammonium bicarbonate in RPLC [33–35] and MA in HILIC [36, 37] can enhance the detection of modified nucleosides. Various mobile phase additives, including formic acid [38], AA [39] (0.1% in RPLC and 0.2% in HILIC), ammonium formate [40] (pH = 3), ammonium acetate [41] (10 mM, pH = 5, 7, and pH 9.0), and ammonium bicarbonate (2 mM and 10 mM) with or without MA, were evaluated, with details of the comparison experiment provided in Supplementary Fig. S2. MS signals of all tested nucleosides under different conditions were summed for comparison. The RPLC condition employed mobile phase A (ddH₂O with 0.1% AA with 50 μM MA) and mobile phase B (methanol with 0.1% AA and 50 μM MA). The gradient program was as follows: 0–1 min (0%–2.5% B), 1–5 min (2.5%–4% B), 5–10 min (4%–5% B), 10–20 min (5%–15% B), 20–21 min (15%–80% B), followed by a 4 min wash at 80% B and a 5 min equilibration at 0% B. The HILIC condition used mobile phase A (95% ACN with 0.2% AA and 50 μM MA) and mobile phase B (50% ACN with 0.2% AA and 50 μM MA). The gradient program was: 0–10 min (0%–1% B), 10–15 min (1%–10% B), 15–20 min (10%–30% B), 20–21 min (30%–50% B), followed by a 4 min wash at 80% B and a 5 min equilibration at 0% B.

Mass spectrometric conditions

Several parameters were evaluated on the 6500 + QTRAP mass spectrometer to optimize sensitivity and minimize DNA and RNA modification fragmentation within the ion source. Design of Experiment approach with Minitab software [Minitab, LLC, Sate College, PA, USA, (2019)] was used, employing a Box-Behnken design to identify key variables and their interactions [42]. The analysis revealed that ion spray voltage (IS), ion source temperature (TEM), collision energy (CE), and declustering potential (DP) significantly influenced the MS response of nucleosides, while other parameters had relatively minor effects (Supplementary Fig. S3A). Among these, the IS was set to 5500 V for both liquid chromatography modes. Detailed adjustments for other parameters are shown in Supplementary Fig. S3B–E. The final MS parameter settings were as follows: for RPLC, the TEM was 400°C, CE 20 eV, DP 50 V, IS 5.5 kV, curtain gas (CUR) 25 psi, the gas 1 (GS1) 55 psi, and gas 2 (GS2) 59 psi; for HILIC, the TEM was 350°C, CE 20 eV, DP 50 V, IS 5.5 kV, CUR 31 psi, GS1 50 psi, and GS2 47 psi. Remaining parameters were set to the instrument's default values. Hydrolytes were analyzed in positive ion mode with an electrospray ion source (ESI).

Nucleic acid modification analysis

To minimize potential contamination of the ESI source, the eluent from the first minute of each LC run was diverted to waste. Commons salts, isotopomers, known artifacts, and potential background noise were systematically excluded to reduce analytical interference. An analysis of monoisotopic masses ($MW < 400$ Da) from curated DNA and RNA modification databases [1, 43, 44] revealed that 99.9% of their decimal mass fractions follow a normal distribution centered at 0.1 ± 0.1 Da (Supplementary Fig. S3F). For deoxyribonucleoside detection, an MRM method was employed, covering precursor $\rightarrow$ product ion transitions ranging from m/z 226.1 $\rightarrow$ 110.1 to m/z 380.1 $\rightarrow$ 264.1, with a 1.0 Da step size across an approximate span of 150 Da. Each sample was analyzed in two injections to increase coverage. The first injection included transitions from m/z 226.1 $\rightarrow$ 110.1 to m/z 305.1 $\rightarrow$ 189.1 and m/z 231.1 $\rightarrow$ 110.1 to m/z 310.1 $\rightarrow$ 189.1. The second injection covered m/z 306.1 $\rightarrow$ 190.1 to m/z 380.1 $\rightarrow$ 264.1 and m/z 311.1 $\rightarrow$ 190.1 to m/z 385.1 $\rightarrow$ 264.1. A dwell time of 5 ms was applied per transition. To focus on rare or modified species, transitions corresponding to canonical deoxyribonucleosides—specifically, m/z 228 or 233 $\rightarrow$ 112 ([dC or (dC + 5) + H]⁺), m/z 243 or 248 $\rightarrow$ 127 ([dT or (dT + 5) + H]⁺), m/z 252 or 257 $\rightarrow$ 136 ([dA or (dA + 5) + H]⁺), and m/z 268 or 273 $\rightarrow$ 152 ([dG or (dG + 5) + H]⁺)—were excluded from the analysis.

A similar MRM strategy was applied to ribonucleosides, with transitions spanning from m/z 242.1 $\rightarrow$ 110.1 to m/z 396.1 $\rightarrow$ 264.1. The first injection targeted m/z 242.1 $\rightarrow$ 110.1 to m/z 321.1 $\rightarrow$ 189.1 and m/z 247.1 $\rightarrow$ 110.1 to m/z 326.1 $\rightarrow$ 189.1, the second covered m/z 322.1 $\rightarrow$ 190.1 to m/z 396.1 $\rightarrow$ 264.1 and m/z 327.1 $\rightarrow$ 190.1 to m/z 401.1 $\rightarrow$ 264.1. Transitions associated with canonical ribonucleosides— m/z 244 or 249 $\rightarrow$ 112 ([rC or (rC + 5) + H]⁺), m/z 245 or 250 $\rightarrow$ 131 ([rU or (rU + 5) + H]⁺), m/z 268 or 273 $\rightarrow$ 136 ([rA or (rA + 5) + H]⁺), and m/z 284 or 289 $\rightarrow$ 152 ([rG or (rG + 5) + H]⁺)—were similarly excluded.

Authentic signals were defined by the concurrent detection of isotopologue pairs: either $[M + H]^+ \rightarrow [M + H - 116]^+$ and $[M + 5 + H]^+ \rightarrow [M + 5 + H - 121]^+$, or $[M + H]^+ \rightarrow [M + H - 132]^+$ and $[M + 5 + H]^+ \rightarrow [M + 5 + H - 137]^+$, with each transition exhibiting a signal-to-noise (S/N) ratio > 2 and an intensity ratio between heavy and light forms ranging from 0.5 to 2. This isotopic criterion ensured confident identification of modified nucleosides while eliminating spurious or artifactual peaks.

LC–HR–MS/MS analysis using SWATH and IDA (DDA–MS²)

High-resolution, nontargeted profiling of deoxynucleosides was performed using a workflow integrating SWATH acquisition [45] and IDA/DDA–MS² [46, 47]. Analyses were conducted on a SCIEX TripleTOF 6600 + mass spectrometer coupled to a Shimadzu LC-40 UHPLC system. Chromatographic separation was achieved on a Waters ACQUITY HSS T3 column (1.8 μ m, 2.1 $\times$ 100 mm) maintained at 40°C with a flow rate of 0.2 ml min^{−1}. Mobile phase A consisted of ddH₂O with 0.1% AA and 50 μ M MA, and mobile phase B consisted of methanol with the same additives. The gradient was as follows: 0–1 min, 0%–2.5% B; 1–5 min, 2.5%–4% B; 5–10 min, 4%–5% B; 10–20 min, 5%–15% B; 20–21 min, 15%–80% B; followed by a 4-min wash at 80% B and a 4-min

re-equilibration. Ionization was performed on an OptiFlow® Turbo V ESI source operated in positive mode with settings slightly modified from those recommended for the 6500 + platform: spray voltage 5500 V, source temperature 400°C, CUR 25 psi, GS1 55 psi, and GS2 60 psi.

For SWATH acquisition, 12 variable windows were generated using the SWATH® Variable Window Calculator (v1.1, SCIEX), applying a 1 m/z overlap at the low-mass edge. MS¹ spectra were collected over m/z 200–500 with a 50 ms accumulation time, a DP of 50 V, and a CE of 5 eV to minimize in-source fragmentation. SWATH–MS² spectra were acquired in high-sensitivity mode over m/z 50–500 using 120 ms per window with DP = 50 V, CE = 15 eV, and a CE spread of 5 eV, resulting in a total cycle time of 1.1 s. SWATH data were processed using NumoFinder, a search engine that incorporates known nucleic acid modifications curated in the Modomics and DNAmoD databases [45], as well as SCIEX OS 2.0 with the Fragment and Neutral Loss Filter plug-in, in which targeted extraction of the characteristic deoxyribose neutral loss was performed using a mass window of 116.0473 ± 0.005 Da. For NumoFinder-based analysis, data interrogation was constrained to the characteristic deoxyribose neutral loss (116.04734 Da), with MS¹ and MS² mass tolerances set to 10 ppm and all remaining parameters kept at default values.

For IDA (DDA–MS²), MS¹ scans were acquired over m/z 200–500 in high-sensitivity mode with a 200 ms accumulation time. The top 10 precursors above 100 cps were isolated for fragmentation, with 50 ms per MS² event, yielding a total cycle time of 0.83 s. To mitigate the inherent bias of the Top-10 DDA strategy toward high-abundance ions and to prevent low-abundance modified nucleosides from being masked by background signals, an inclusion list was implemented. Specifically, the m/z values corresponding to 5mC, 5hmC, 5fC, and 5caC were added to the inclusion list to force MS² acquisition of these species regardless of their MS¹ intensities. IDA datasets were analyzed with SCIEX OS 2.0 using the Fragment and Neutral Loss Filter plug-in.

FA treatment of genomic DNA

A FA solution was prepared by diluting a 37% (w/v) aqueous FA solution with phosphate buffered saline (PBS). Isotope-labeled genomic DNA isolated from mES cells was used in the reaction. In a 100 μ l reaction, 100 μ g of DNA was incubated with 25 mM FA at 37°C for 3 h. For the control DNA, the FA solution was replaced with an equal volume of 1 $\times$ PBS. Following the reaction, sodium acetate was added to a final concentration of 300 mM, and the mixture was thoroughly mixed. An equal volume of isopropanol was then added to precipitate the DNA. The resulting DNA pellet was dissolved in 100 μ l of ddH₂O for subsequent analysis.

Preparation of N⁴-hydroxymethyl-deoxycytidine standards

Solutions (1 mM) of three deoxynucleosides—deoxycytidine (dC), 5-methyl-deoxycytidine (5mC), and 5-hydroxymethyl-deoxycytidine (5hmC), along with their respective isotopically labeled counterparts (¹⁵N₃-dC, ¹⁵N₂, ¹³C-5mC, and D₃-5hmC) were prepared in 100 mM ammonium acetate buffer (pH 7.0). Each solution was incubated with excess FA (final concentration: 50 mM) at 37°C for 2 h. The reaction progress was monitored by high performance liquid chromatography

with ultraviolet (UV) detection. In each reaction, two major UV-absorbing peaks were observed: the earlier-eluting peak corresponded to the untreated starting nucleoside, while the later-eluting peak represented the N^4 -hydroxymethylated adduct. The identity of each product was confirmed in MRM-IDA-EPI mode.

Low resolution MS for DNA adduct identification

Candidate deoxynucleoside fragments were acquired using an MRM–information-dependent acquisition–enhanced product ion scan (MRM–IDA–EPI) approach. EPI scans were automatically triggered by MRM features that met predefined acquisition criteria. These ions were dynamically added to the inclusion list for fragmentation. Each MRM transition was assigned a dwell time of 5 ms. EPI scans were conducted in the ion trap over an m/z range of 50–400, at a scan rate of 10 000 amu/s, with dynamic filling enabled. The intensity threshold for scan triggering was set at 2000 counts per second (cps), with a mass tolerance of 250 milli Daltons (mDa) and a detection window of 1. All other parameters were consistent with those described previously.

Targeted DNA adducts analysis by HRMS

Targeted analysis of DNA adducts was performed using a Nexera LC-40 system (Shimadzu) coupled to a Q EXactive Plus HRMS (Thermo Fisher Scientific). Chromatographic separation was carried out on an HSS T3 column (1.7 μ m, 2.1 $\times$ 100 mm) maintained at 35°C, with a flow rate of 0.3 ml/min. The mobile phases consisted of (A) 100% water with 0.1% formic acid and (B) methanol with 0.1% formic acid. The gradient program was as follows: 0–1 min, 0% B; 1–1.1 min, linear increase to 5% B; 1.1–2.5 min, hold at 5% B; 2.5–2.6 min, ramp to 80% B; 2.6–3.0 min, wash at 80% B; and 3.0–6.0 min, re-equilibration at 0% B.

Mass spectrometric detection was performed in positive ion mode using a heated electrospray ionization source. Source conditions were as follows: sheath gas flow at 40 arbitrary units, auxiliary gas at 10, spray voltage at +3.5 kV, capillary temperature at 320°C, and auxiliary gas heater temperature at 300°C. An inclusion list was used to monitor targeted DNA adducts. Each injection comprised a full scan acquired concurrently with a parallel reaction monitoring (PRM) scan. Full scan parameters were: m/z range of 50–700, resolution of 45 000 (at m/z 200), automatic gain control (AGC) target of 3×10^6 , and maximum injection time (IT) of 100 ms. PRM settings included a resolution of 17 500, AGC target of 1×10^5 , maximum IT of 50 ms, loop count of 6, and an isolation window of 1 m/z .

Results

Dual neutral loss scanning with equimolar ^{13}C - and ^{12}C -incorporated (deoxy)ribose enables sensitive detection of low-abundance nucleosides

Uniformly-labeled [$\text{U-}^{13}\text{C}$]glucose has been used to trace RNA modifications in bacteria and yeast. In these organism, full ^{13}C incorporation into both ribose and nucleobase moieties can be achieved by culturing cells over many generations in minimal medium where glucose is the sole carbon source [48, 49]. In mammalian cells, however, the situation is more complex. Ribose and deoxyribose sugars are synthesized directly through the pentose phosphate pathway (PPP), a key route

downstream of glucose metabolism [50]. By contrast, nucleobase synthesis relies on multiple precursors—such as glutamine, asparagine, and glycine—that are themselves derived from glucose only after additional metabolic conversions [51, 52]. As a result, the ribose moiety incorporates the ^{13}C label more rapidly and efficiently than the nucleobase. Achieving uniform nucleobase labeling in mammalian cells using [$\text{U-}^{13}\text{C}$]glucose is essential to avoid spectral complexity, as partial labeling broadens the isotopic distribution of fragment ions and reduces sensitivity for detecting low-abundance modifications. However, attaining this level of uniform labeling typically requires prolonged incubation times, which may increase nucleobase recycling—in which RNA-derived modifications can appear in DNA analyses, potentially leading to false positives [51, 53, 54]. To address these challenges, we decided to selectively label a subset of the sugar moieties in cellular DNA and RNA while minimizing nucleobase incorporation. This approach is expected to preserve strong, well-defined isotopologue signals that can be confidently detected and distinguished in the mass spectrometer—an especially critical consideration for rare or low-abundance nucleoside modifications. When cells are cultured in medium containing [$\text{U-}^{13}\text{C}$]glucose, the labeled carbon atoms are incorporated into pentose sugars, producing heavy-labeled analogs of their native ^{12}C forms (Fig. 1A). This metabolic labeling strategy gives rise to two characteristic neutral loss fragments: 116 Da and 121 Da for deoxyribose, and 132 Da and 137 Da for ribose. In the chromatogram, these pairs produce distinct MRM transitions with identical retention times—specifically, $[\text{M} + \text{H}]^+ \rightarrow [\text{M} - 116 + \text{H}]^+$ and $[\text{M} + 5 + \text{H}]^+ \rightarrow [\text{M} + 5 - 121 + \text{H}]^+$ for deoxyribonucleosides and $[\text{M} + \text{H}]^+ \rightarrow [\text{M} - 132 + \text{H}]^+$ and $[\text{M} + 5 + \text{H}]^+ \rightarrow [\text{M} + 5 - 137 + \text{H}]^+$ for ribonucleosides.

To achieve preferential labeling of sugar moieties while minimizing isotopic incorporation into nucleobases, we optimized the metabolic labeling conditions in mES cells. Culturing the cells for 48 h in medium supplemented with 20 mM [$\text{U-}^{13}\text{C}$]glucose and 5 mM [$\text{U-}^{12}\text{C}$]glucose resulted in the desired partial labeling of the deoxyribose moiety in genomic nucleosides. MS analysis revealed a distinct isotopic distribution in which the unlabeled ($^{13}\text{C}_0$) and fully labeled ($^{13}\text{C}_5$) deoxyribonucleosides were present at nearly equal abundance, each comprising 33%–42% of the total deoxyribose (Fig. 1B). All other partially labeled species containing 1–4 ^{13}C atoms in the sugar ring accounted for 19%–27%. This labeling pattern was consistent across all four canonical deoxyribonucleosides (dA, dT, dC, and dG). Notably, a small but consistent +1 isotopologue (~4%) was also observed in cells cultured with unlabeled glucose, and its abundance remained largely unchanged under [$\text{U-}^{13}\text{C}$]glucose labeling (data not shown). This suggests that the +1 species may arise from natural isotopic background. In contrast, the +2 to +4 labeled fractions increased substantially with ^{13}C glucose, likely reflecting carbon rearrangements via the nonoxidative PPP and the incorporation of unlabeled carbon intermediates through gluconeogenesis or other metabolic fluxes [55]. A comparable 1:1 sugar labeling ratio ($^{13}\text{C}_5$ versus $^{12}\text{C}_5$) was also achieved in ribonucleosides from cellular RNA (Supplementary Fig. S4A). Importantly, under these conditions, ^{13}C incorporation into nucleobases was largely suppressed. Despite the baseline presence of natural isotopic variants (ranging from 7% to 17%), over 80% of nucleobases—excluding naturally occurring isotopes—remained unlabeled

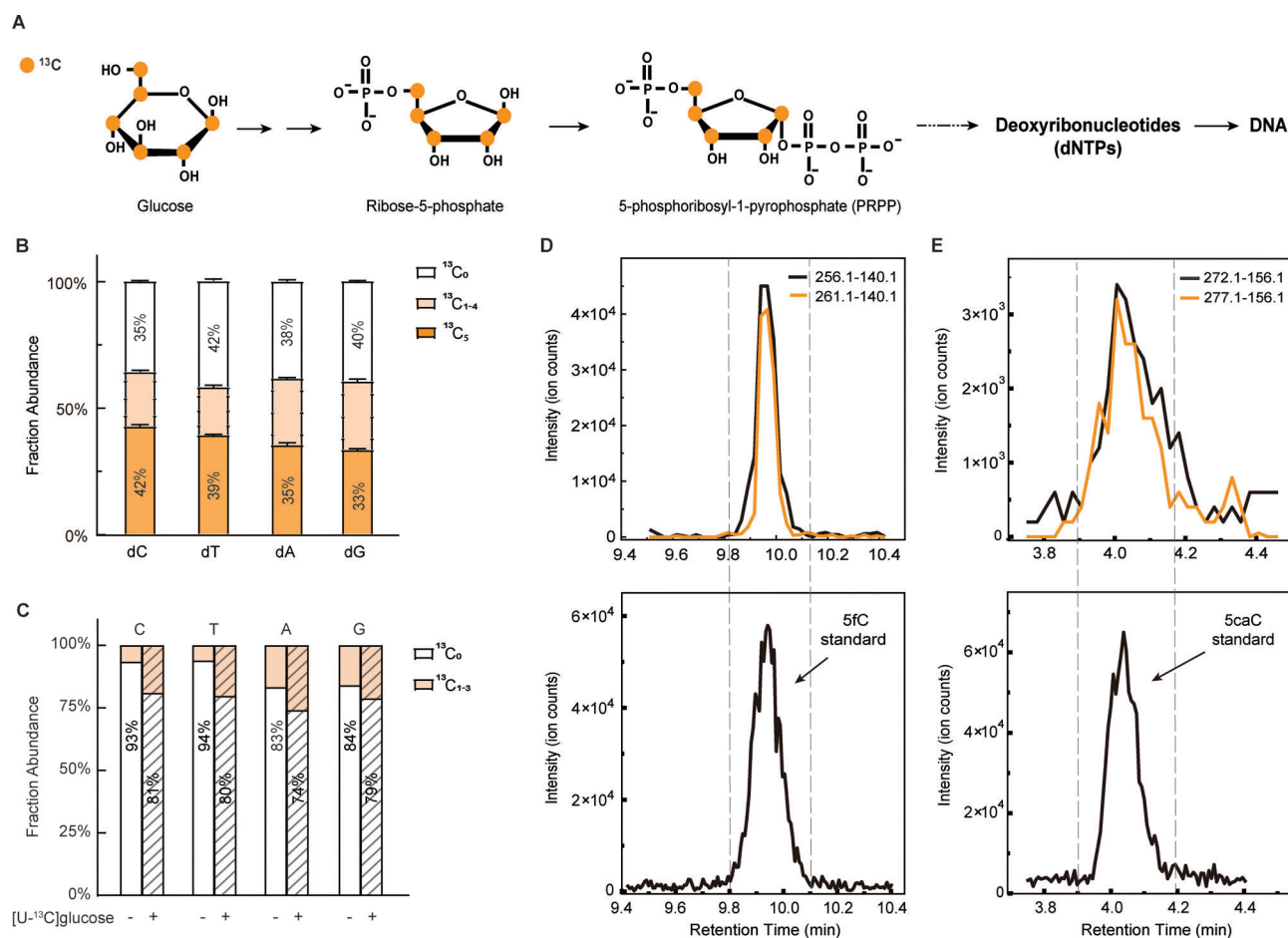


Figure 1. Metabolic labeling of cellular DNA with [U- ^{13}C]glucose for high-sensitivity detection of modified nucleosides. **(A)** Schematic of ^{13}C incorporation into the ribose moiety of nucleotides via cellular metabolism of [U- ^{13}C]glucose. The generated ^{13}C -labeled pentose sugars are incorporated into PRPP, which serves as a precursor for dNTP biosynthesis. Orange dots denote ^{13}C atoms. **(B)** Approximate 1:1 ratio of fully ^{13}C -labeled ($^{13}\text{C}_5$, orange) and unlabeled ($^{13}\text{C}_0$, unfilled) deoxyribose species in genomic DNA from mES cells fed with a mix of [U- ^{13}C]glucose and [U- ^{12}C]glucose. Bar plots depict the distribution of deoxyribonucleosides (dC, dT, dA, dG) with sugar moieties containing 0, 1–4, or 5 ^{13}C atoms, representing unlabeled (unfilled), partially labeled (light orange), and fully labeled (orange) species, respectively. Only the proportions of the $^{13}\text{C}_0$ and $^{13}\text{C}_5$ fractions are shown, as these are essential for downstream isotopic pairing analysis. Data represent the mean of three independent biological replicates, with error bars indicating standard error of the mean. **(C)** DNA nucleobases in mES cells remained largely unlabeled upon feeding with [U- ^{13}C]glucose. Unfilled bars represent the natural isotopic distribution of nucleobases in dN from cells cultured with [U- ^{12}C]glucose, while hatched bars show the isotopologue distribution of nucleobases in $^{13}\text{C}_5$ -labeled dN from [U- ^{13}C]glucose-cultured cells. Bar heights indicate relative abundance. The increase in light orange bars under [U- ^{13}C]glucose conditions reflects ^{13}C incorporation into the nucleobase portion of $^{13}\text{C}_5$ -labeled dN. **(D, E)** High-sensitivity detection of 5fC **(D)** and 5caC **(E)** nucleosides in genomic DNA from mES cells. Overlaid extracted ion chromatograms (EIC) from RPLC analysis display the $^{13}\text{C}_5$ -labeled (orange) and unlabeled (black) species. Ion transitions (precursor → fragment) are displayed in the top-right corner. Peaks from authentic standards (blue) are shown below as reference. Genomic DNA was isolated from WT mES cells cultured for 2 days in medium containing 20 mM [U- ^{13}C]glucose and 5 mM [U- ^{12}C]glucose.

(Fig. 1C and Supplementary Fig. S4B). We further validated this strategy in HEK293T cells (Supplementary Fig. S4C), demonstrating that this metabolic labeling approach is robust and adaptable with proper culture adjustments for proliferative cell lines. This predominant sugar labeling, together with balanced peak intensities between light- and heavy-labeled deoxyribonucleosides, facilitates parallel detecting of both isotopologue species.

To evaluate the sensitivity of this ribose-favored labeling strategy in the context of MRM screening, we focused on two rare DNA modifications—5fC and 5caC—which occur at an estimated frequency of 1 in 10^6 – 10^7 nucleotides in mES cells [20]. While 5fC has been occasionally reported in untargeted MS workflows [29], reliable detection of these rare modifications—especially 5caC—remains limited. Using MRM, we set the Q1 isolation window to 1.0 Da and scanned

precursor ions from m/z 226.1 to m/z 380.1 in 1.0 Da increments. Product ions ranging from m/z 110.1 to m/z 264.1 were monitored in Q3, covering a total detection range of ~ 150 Da. Common nucleosides (A, T, C, G) were excluded from data collection to reduce background. With 10 μg of input DNA per injection, QqQ-MS was able to screen this range at unit resolution, specifically targeting neutral losses of 116 and 121 Da. This approach successfully captured distinct ion pairs corresponding to 5fC and 5caC, with light and $^{13}\text{C}_5$ -labeled forms co-eluting at identical retention times and displaying comparable intensities (Fig. 1D and E). These results demonstrate that our (deoxy)ribose-labeling strategy enables sensitive, untargeted detection of rare nucleosides at concentrations as low as 1 ppm in MRM mode. The method is broadly applicable for discovering novel or trace DNA modifications.

To assess the performance of our metabolic-labeling workflow on HRMS platforms, we profiled DNA modifications on a TripleTOF 6600+, using both SWATH acquisition and DDA-MS². Using the NumoFinder search library, a panel of known cytosine modifications, including 5fC and 5caC, was readily identified in mixed deoxynucleoside standards, demonstrating that the spectral library and analysis workflow were capable of detecting these modifications under controlled conditions. In contrast, when applied to genomic DNA isolated from mES cells, only the high-abundance modifications 5mC and 5hmC were consistently detected, whereas 5fC and 5caC were not observed. Consistently, during DDA-MS² data processing, even when the mass tolerance for neutral-loss searching was relaxed to 50 ppm, 5fC and 5caC remained undetectable. To exclude the possibility of false negatives, we further expanded the retention time window to ± 2 min and relaxed the mass tolerance to ± 0.05 Da. Even under these permissive conditions, no putative signals corresponding to 5fC or 5caC were observed, whereas only the high-abundance modifications 5mC and 5hmC were detectable (Supplementary Fig. S5). Published high-resolution datasets, including S462 cells analyzed on a ZenoTOF 7600+, similarly reported no detectable 5fC or 5caC despite substantial 5hmC levels [45]. Collectively, these results demonstrate that MRM offers superior sensitivity for unbiased pre-screening of extremely low-abundance nucleosides, and therefore we conducted subsequent discovery-oriented screening on this platform.

¹³C-(deoxy)ribose labeling enables discrimination of exogenous nucleosides

Commercially available enzymes, such as nucleases and phosphatases used for DNA and RNA digestion, are often derived from bacterial sources. As a result, noneukaryotic nucleoside species from bacterial genomes may be inadvertently introduced during sample preparation, potentially confounding LC-MS/MS analyses [56]. For instance, nucleosides such as 5mC, 3-methylcytosine, 4mC, and 6mA are frequently detected in enzymatic digests, yet such modified species cannot conclusively establish their presence in eukaryotic genomes [8]. To address this issue, we leveraged ¹³C-labeling of (deoxy)ribose through the nucleotide synthesis pathway in mammalian cells, stabilizing the ¹²C/¹³C ratio of canonical nucleosides at ~ 1 . This internal reference ensures that bona fide endogenous modifications exhibit the same isotopic ratio. In contrast, modifications deviating from this ratio are likely artifacts introduced post-lysis, such as contaminants from enzymatic reagents. This strategy substantially improves the fidelity of modification detection by distinguishing genuine cellular nucleosides from exogenous contaminants.

To validate this approach, we cultured mES cells lacking DNMTs *Dnmt1*, *Dnmt3a*, and *Dnmt3b* (*Dnmt* TKO) in [U-¹³C]glucose and performed MRM scans to construct global modification profiles. After removing isotope-related and metal ion-associated species, we identified 130 and 196 transitions with neutral loss of 116 Da on RPLC and HILIC columns, respectively (Supplementary Tables S1 and S2). However, many transitions lacked corresponding ¹³C-labeled peaks or showed significantly skewed labeling ratios (>1.5 or <0.5), suggesting they were false positives arising from nondeoxynucleosides or noneukaryotic sources (Fig. 2A and B). For example, although *Dnmt* TKO cells are expected to

lack genomic 5mC, we detected residual unlabeled 5mC peak at ~ 1 per 10^6 dC. Notably, no matching ¹³C-labeled 5mC peak was observed at the same retention time, indicating that the 5mC peak is artifact, likely introduced via commercial reagents (Fig. 2C) [56]. After rigorous scrutiny to eliminate such artifacts, only 23 and 41 putative modified deoxynucleosides were retained as true positives on the RPLC and HILIC columns, respectively (see the final columns of Supplementary Tables S1 and S2; see Supplementary Table S3 for the overlap between RPLC and HILIC analyses), yielding feature sets that are highly enriched for genuine endogenous nucleoside modifications. The false-positive issue originates primarily from non-nucleoside species or exogenous nucleosides introduced during sample collection and handling and are not unique to triple quadrupole-based workflows. Indeed, we observed comparable examples on the HRMS platform (Supplementary Fig. S6). For instance, ion pairs such as m/z 221.1159 $\rightarrow$ 105.0685 and 315.1543 $\rightarrow$ 199.1070 show a neutral loss of 116.0473 Da consistent with a deoxyribose moiety; however, the absence of the corresponding +5 Da isotopically labeled counterparts unambiguously indicates that these species are exogenous contaminants or non-nucleoside molecules. Because the source of such signals lies upstream of MS acquisition, their occurrence is independent of the mass spectrometric resolution applied.

One particularly debated DNA modification is 6mA, whose reported levels in mammalian cells span a wide range—from 1 to 8000 ppm—across various cell types and developmental stages, including mES cells and pig embryos [14–16]. Some studies have proposed that the detected 6mA may actually stem from misincorporation of ribonucleoside into DNA [51, 53, 54]. However, our analysis did not support the existence of endogenous 6mA in the mES genome, as suggested by the absence of ¹³C₅-labeled 6mA peak (Fig. 2D). Although the labeling strategy does not capture misincorporation events directly, our findings are consistent with recent studies that question the existence of 6mA in mammalian DNA [19, 20, 57]. Importantly, the peak intensity of the unlabeled 6mA peak was comparable to that of authentic 5fC in WT cells, emphasizing the risk of misinterpreting such artifacts as genuine epigenetic marks. Our approach therefore addresses a key limitation of enzyme-based assays, which cannot definitively determine the cellular origin of 6mA [56]. Even under untargeted conditions, where sensitivity is inherently reduced, our method reliably confirms the absence of endogenous 6mA at ppm levels, aligning with prior HRMS studies [20, 57].

A similar pattern was observed in RNA modification profiling, where nearly half of the species failed to meet the criteria of co-detection of both unlabeled and labeled peaks (Supplementary Fig. S7A and B, and Supplementary Tables S4 and S5). Notably, certain RNA modifications showed column-specific detectability—for example, *N*⁷-methylguanosine (*m*⁷G) was observed only on the RPLC column, while ribosylinosine (rI) was exclusively detected on the HILIC column (Supplementary Fig. S7C and D). This highlights the complementary strengths of the two separation methods for achieving comprehensive modification coverage [58].

In summary, our (deoxy)ribose-labeling strategy enables clear discrimination between endogenous and exogenous nucleosides, significantly reducing false positives and allowing for untargeted detection of authentic DNA and RNA modifications with high sensitivity and reliability.

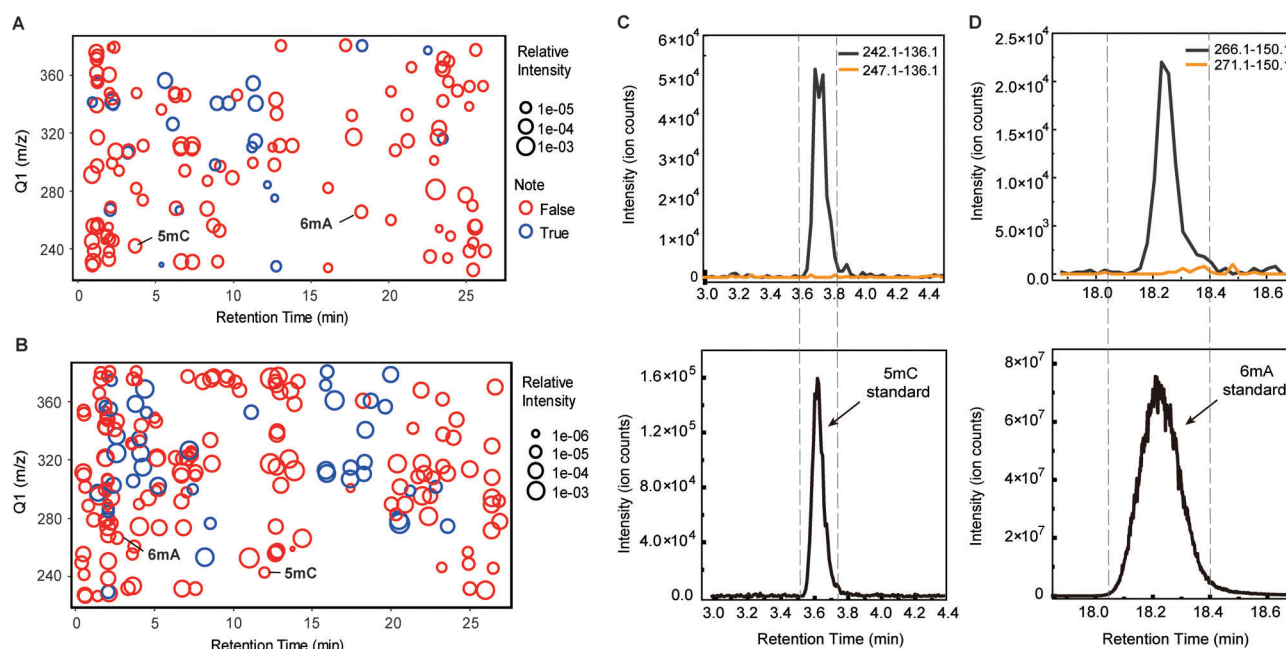


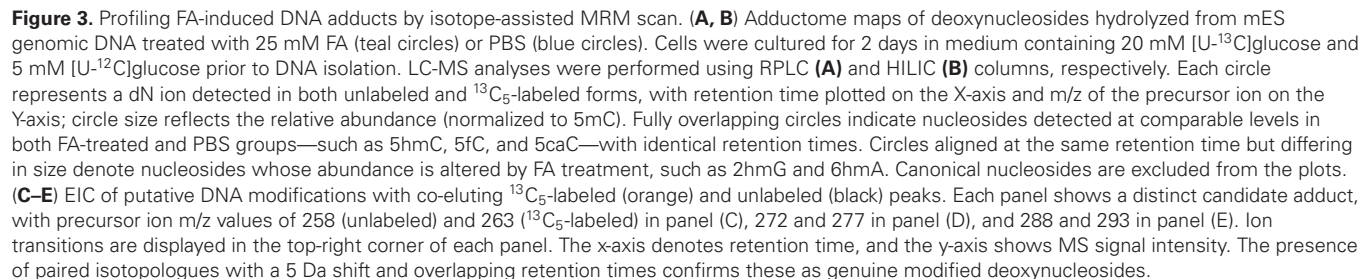
Figure 2. ^{13}C -labeling of deoxyribose enables discrimination of false-positive species in DNA modification analysis. (A, B) Filtering of candidate noncanonical nucleoside hits (blue) from mouse genomic DNA to remove the false positives (red). MRM transition maps show precursor-to-fragment ion pairs with a characteristic neutral loss of 116 Da. Genomic DNA was extracted from *Dnmt* triple knockout (TKO) mES cells, analyzed using either RPLC (A) or HILIC (B) columns. Cells were cultured for 2 days in medium containing 20 mM [U - ^{13}C]glucose and 5mM [U - ^{13}C]glucose. Data are presented as bubble plots, with the x-axis representing the mass-to-charge ratio (m/z), the y-axis denoting retention time, and the size of each circle corresponding to the ion intensity relative to deoxyadenosine +2 (254.1 $\rightarrow$ 138.1) to correct for injection variability. The four canonical DNA nucleosides are excluded for clarity. Blue circles represent dN ions detected in both unlabeled and $^{13}C_5$ -labeled forms (classified as true), while red circles indicate ions observed only in the unlabeled form (classified as false). Data are representative of three independent experiments. (C, D) Identification of false-positive calls exemplified by 5mC (C) and 6mA (D) in genomic DNA from *Dnmt* TKO mES cells. Overlaid EIC from RPLC analysis for $^{13}C_5$ -labeled (orange) and unlabeled (black) species. Peaks from authentic standards are shown below as reference. Ion transitions are shown at the top right of each panel. The y-axis represents MS signal intensity, and the x-axis shows retention time.

Dual neutral loss scan reveals novel adducts in FA-treated genomic DNA

Given that unit-resolution mass analyzers such as quadrupoles and ion traps can only resolve mass differences of 0.5–1 Da, loss of 116 Da or 132 Da alone is insufficient to unambiguously assign the compound as a deoxyribonucleoside or ribonucleoside. To overcome this limitation, we incorporated stable isotope labeling, such that true sugar-containing species would produce paired neutral losses—116/121 Da for deoxyribose and 132/137 Da for ribose—corresponding to unlabeled and $^{13}C_5$ -labeled forms, respectively. To support this interpretation, we consulted the Beynon table and applied chemical valence rules to infer plausible molecular formulas. The observed neutral losses of 116 and 121 Da are consistent with $C_5H_8O_3$ (deoxyribose) and $C_5H_{12}N_2O$, respectively, while the 132 and 137 Da losses match formulas such as $C_5H_8O_4$ (ribose), $C_5H_{12}N_2O_2$, and $C_5H_{16}N_4$. While some of these formulas match known nitrogen-containing metabolites (e.g. creatinol, arginine acid, spermidine), their biosynthesis involves nitrogen–carbon rearrangements and degradation of the glucose backbone, which substantially limits the likelihood of full $^{13}C_5$ incorporation within the short labeling window of our experimental design. Therefore, the most parsimonious and biochemically plausible interpretation is that the paired neutral losses of 116/121 Da and 132/137 Da reflect light and heavy deoxyribose and ribose moieties, respectively. Although QQQ-MS lacks the mass accuracy and structural resolution of HRMS, our QQQ-based paired neutral loss MS workflow provides a high-sensitivity, isotope-resolved screen-

ing platform. By selectively detecting molecules that show both light and heavy sugar losses, it enables confident assignment of (deoxy)ribose-containing candidates, which can then be structurally characterized using HRMS or nuclear magnetic resonance (NMR).

To demonstrate the utility of this isotope-labeling strategy in identifying both known and novel nucleoside modifications, we applied our method to investigate FA-induced DNA adducts. FA, a ubiquitous reactive metabolite generated during cellular metabolism, is known for its electrophilic reactivity and propensity to form DNA adducts, particularly N^2 -hydroxymethylguanine (2hmG) and N^6 -hydroxymethyladenine (6hmA) *in vivo* [59, 60]. *In vitro* studies have revealed a broader array of adducts, prompting the development of diverse MS-based approaches—including constant neutral loss (CNL), data-dependent, CNL triggered MS³ (DDA-CNLS-MS³), and DDA-CNLS-MS²—to characterize FA-induced lesions [46]. Using our dual-isotope labeling strategy, we analyzed mES genomic DNA treated with FA or PBS. A total of 40 and 60 ion pairs were identified in RPLC and HILIC modes, respectively, from the PBS group, with an additional 45 and 60 ions detected in the FA-treated group (Fig. 3A and B, and Supplementary Table S6). Cross-referencing with previously reported adducts confirmed that our method reliably captured all known monoadducts detected by earlier HRMS-based approaches [46], and crucially, also uncovered several novel candidates. Among these were putative structures such as 3-hydroxymethyl-thymine (m/z 273; Supplementary Fig. S8A), N^4 -ethoxymethyl-cytosine



Interestingly, while most previously reported monoadducts predominantly occur on adenine bases [46], our analysis revealed multiple adducts on cytosine and its epigenetically modified derivatives. These included N^4 -hydroxymethylcytosine (4hmC, m/z 258), N^4 -hydroxymethyl-5mC (4hm5mC, m/z 272), and N^4 -hydroxymethyl-5hmC (dihmC, m/z 288) (Fig. 3C–E). To confirm the identities of these adducts, we synthesized in-house standards by reacting FA with deoxynucleosides dC, 5mdC, and 5hmdC, along with their isotopically labeled counterparts ($^{15}\text{N}_3$ -dC, $^{15}\text{N}_2$, ^{13}C -5mdC, and D_3 -5hmdC) [31, 46]. These standards were analyzed using LC-QqQ-MS, and their identities were validated through EPI-MS² fragmentation spectra (Supplementary Fig. S9A–F). Parallel analysis of FA-treated

These hydroxymethyl adducts are known to be unstable due to the reversible nature of Schiff base formation between FA and the primary amines on nucleobases. Prior studies have shown that the monophosphate form of 4hmdC degrades more rapidly than those of 2hmdG and 6hmdA [31], indicating that 4hmC is inherently unstable and prone to loss during sample processing. This aligns with our observation that 4hmC was 5.5-fold more abundant in RPLC than in

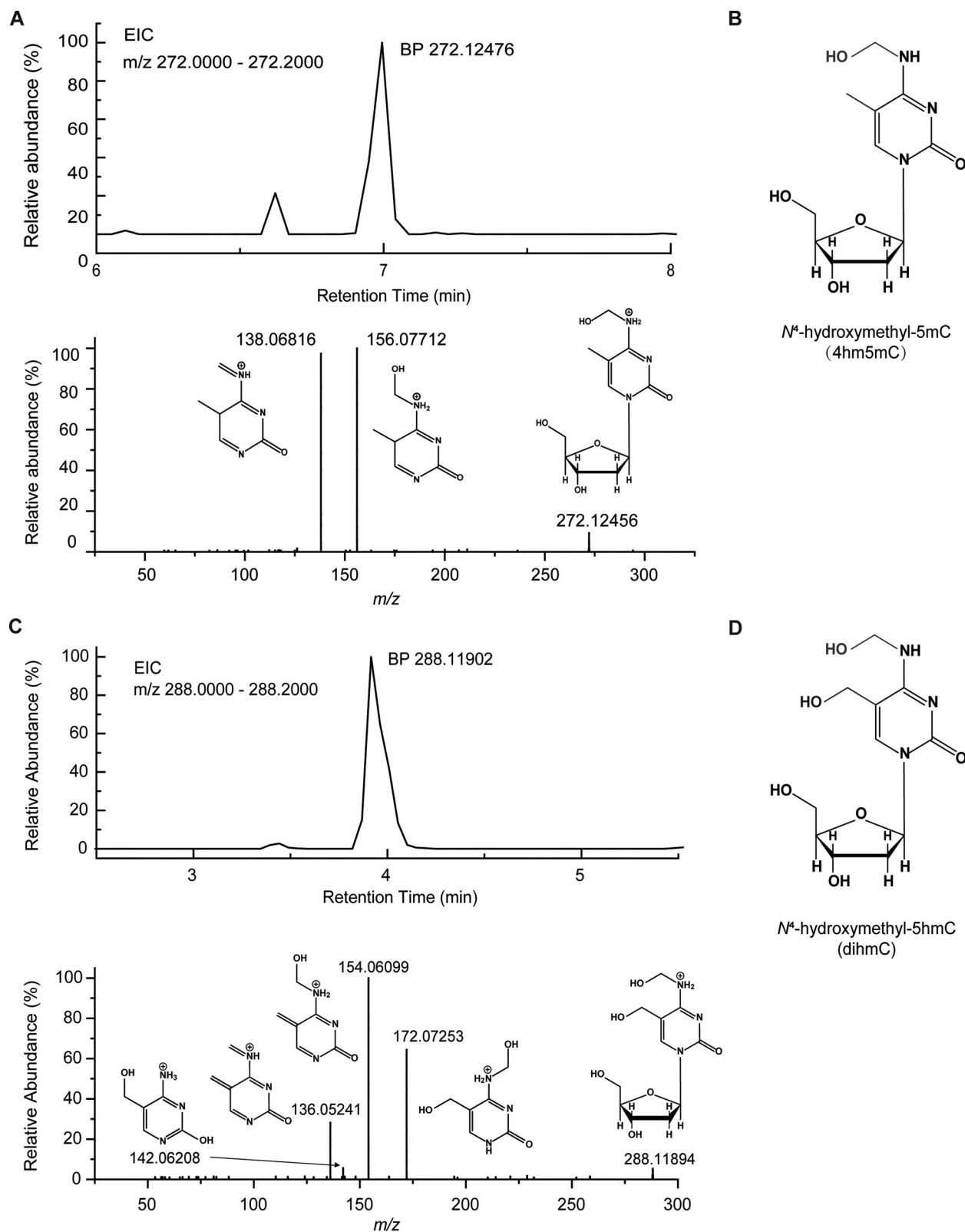


Figure 4. Structural characterization of modified deoxynucleosides from FA-treated genomic DNA with m/z 272 and 288 by high-resolution MS. (**A, B**) EIC (A, top) for ions within the m/z range 272.0000–272.2000, showing a base peak at m/z 272.12476. The corresponding MS² spectrum (A, bottom) displays the fragmentation pattern with proposed structures for each fragment. The inferred chemical structure is illustrated in panel (B). (**C, D**) EIC (C, top) for ions within the m/z range 288.0000–288.2000, highlighting a base peak at m/z 288.11902. The MS² spectrum (C, bottom) shows the fragmentation pattern and corresponding fragment structures, with the deduced structure shown in panel (D). Hydroxymethyl groups located at the N4 position of 5mC and 5hmC are indicated in blue. Proposed chemical structures were deduced based on the accurate masses of characteristic fragment ions. MS was performed using high-resolution PRM mode on hydrolyzed genomic DNA treated with FA.

HILIC mode, presumably due to the reduced sample handling time associated with RPLC (Fig. 3A and B). We hypothesized that 4hm5mC and dihmC might exhibit similar instability and assessed their half-lives using our in-house standards by UPLC. The calculated half-lives were 204.4 min for 4hmC, 16.5 min for 4hm5mC, and 25.1 min for dihmC, respectively (Supplementary Fig. S10A–C), confirming their inherent instability and the technical challenge they pose for detection and quantification.

Our results highlight that even with LRMS, the use of isotopic dual labeling enables the confident recognition of true nucleoside-derived ions—not only for known modifications validated by their respective standards, but also for novel modified species whose nucleoside origin was subsequently confirmed by HRMS. This underscores the potential of our approach as a sensitive initial screening strategy for uncovering previously uncharacterized or labile DNA modifications.

Discussion

The discovery of novel nucleoside modifications offers valuable opportunities to gain insights into fundamental biological processes and disease mechanisms. Existing detection strategies generally fall into two categories: modifying enzyme-guided approaches and untargeted modification profiling. Enzyme-guided methods rely on known enzymatic pathways, often identifying modifications related to established ones due to the conserved nature of enzyme superfamilies across species. In contrast, untargeted profiling offers the potential to uncover entirely new modifications but demands stringent sample preparation and detection protocols to mitigate artifacts and false positives. In this study, we introduced a dual neutral loss strategy that integrates [U-¹³C]glucose metabolic labeling with untargeted MRM-MS. Culturing mouse ES cells with an optimized mixture of [U-¹³C]glucose and natural ¹²C-glucose enabled near 1:1 labeling of the (deoxy)ribose moiety in genomic DNA. This isotopic pairing allowed the differentiation of endogenous nucleosides from contaminants, increasing the confidence in assigning genuine nucleoside features. Streamlined sample preparation further reduced processing-induced artifacts such as oxidation (e.g. 8oxo-dG) and deamination (e.g. dU). Finally, the use of complementary RPLC and HILIC columns improved resolution of co-eluting species, as demonstrated by the selective detection of certain modifications: m7G on the RPLC column and rI on the HILIC column.

Leveraging this newly developed metabolic (deoxy)ribose-labeling approach, we uncovered several previously unreported FA-induced DNA adducts. These include dihmC, whose structure was validated by HRMS, as well as 3-hydroxymethyl-thymine, N⁴-ethoxymethyl-cytosine, N⁶-ethoxymethyl-N⁶-methyl-adenine, and N²-ethoxymethanol-hydroxymethyl-guanine. Notably, while standard base damage often resolves to canonical forms (e.g. 8oxo-dG repairing to dG), damage to epigenetic modifications like 4hm5mC and dihmC may result in epigenetic information loss, potentially leading to more profound biological consequences. Supporting this idea, a recent study demonstrated that alkylation damage to 5mC, forming ε5mC, can inactivate both glycosylases and epigenetic enzymes [61], amplifying the detrimental effect on epigenetic modifications. Additionally, FA exposure has previously been reported to induce epigenetic alterations by inhibiting the activity of S-adenosylmethionine synthase isoform type-1 (MAT1A), thereby limiting the cel-

lular methyl donor pool [62]. In addition to this indirect mechanism, our findings reveal a complementary and potentially more direct pathway: FA can also disrupt the epigenetic landscape by chemically modifying DNA bases that contain epigenetic modification. This direct reactivity with modified nucleobases—such as 5mC and its oxidized derivatives—not only alters their chemical structure but may also interfere with the recognition and processing by reader, writer, and eraser proteins, thereby contributing to aberrant epigenetic regulation following FA exposure.

The MRM-based screening has been widely employed in adduct detection due to its high sensitivity [63, 64], and holds great potential for untargeted discovery of DNA and RNA modifications. Various forms of isotopic labeling have been integrated into MRM workflows. Notably, previous studies have utilized *in vivo* labeling of nucleotides with ¹⁵NH₄Cl or [U-¹³C]glucose in *Escherichia coli* and *Saccharomyces cerevisiae*, enabling the detection of nucleoside modifications by replacing all nitrogen or carbon atoms within the nucleoside backbone [48, 49, 65]. These strategies rely on co-analysis of labeled and unlabeled samples, where candidate modifications are identified based on co-elution and characteristic mass differences of the labeled isotopomers. Such approaches have opened new avenues for identifying previously uncharacterized nucleoside structures. Inspired by these microbial labeling strategies, we adapted and simplified this concept for application in mammalian cells. Specifically, we mixed the regular glucose in the culture medium with uniformly labeled [U-¹³C]glucose, enabling the incorporation of ¹³C into the (deoxy)ribose moiety of endogenous nucleosides at a 1:1 ratio of ¹³C₅ to ¹²C₅. This approach achieves selective labeling without perturbing global cellular metabolism and is particularly compatible with mammalian cells, where full nucleotide replacement is often impractical. Because only the ribose moiety is labeled, the resulting isotopic signature remains simple and avoids the analytical complexity often introduced by base labeling. By focusing on the (deoxy)ribose unit as a labeling handle, our method facilitates the discrimination of endogenous nucleosides from exogenous artifacts through isotopic pairing. Importantly, this authenticity check is achieved within a single sample, eliminating the need for separate labeled and unlabeled controls.

Our approach has several limitations. First, while the glucose labeling strategy is generally applicable to mammalian cells with active pentose phosphate metabolism, its efficiency may be limited in organisms with reduced or alternative ribose synthesis pathways, such as certain obligate parasites or bacteria lacking key pentose phosphate pathway enzymes. For such cells, labeling strategies must be tailored to their specific metabolic pathways. The labeling duration and concentration of [U-¹³C]glucose used for labeling also depend on specific cell line and cell culture conditions. Second, effective DNA labeling depends on DNA replication; therefore, cells that are not actively undergoing the cell cycle, such as terminally differentiated tissue cells, cannot be efficiently labeled. Furthermore, the MRM-MS method primarily analyzes base modifications following the loss of neutral ribose, limiting the detection of modifications on the sugar moiety. Thus, exploring modifications on the deoxyribose moiety in nucleic acids will require the development of alternative labeling strategies and MS configurations.

In conclusion, novel DNA modifications likely occur at low abundance, and existing detection workflows face challenges

related to specificity, sensitivity, and processing artifacts. By combining the high sensitivity of MRM-MS with the specificity provided by (deoxy)ribose-isotope labeling and streamlined sample preparation, we present a robust and efficient platform for untargeted profiling of DNA and RNA modifications. This approach offers a promising framework for discovering novel nucleoside modifications.

Acknowledgements

We acknowledge Hailin Wang, Kanglin Zhang, Bifeng Yuan, Saulius Klimasauskas, Jing Xu, and Zhidan Liang for critical reading of the manuscript; the Core Facility of Shanghai Medical College, Fudan University, for their technical supports.

Author contributions: Shao-Qin Rong (Conceptualization [lead], Data curation [lead], Formal analysis [lead], Methodology [lead]), Zhi-Han Jin (Data curation [equal], Formal analysis [equal]), Dong-Rui Yin (Data curation [equal], Formal analysis [equal]), Jun Yao (Formal analysis [supporting], Resources [supporting]), Guoliang Xu (Funding acquisition [equal], Writing – review & editing [supporting]), Ping Zhu (Funding acquisition [supporting]), Hai Gao (Investigation [equal], Project administration [equal]), Dan Zhou (Funding acquisition [lead], Investigation [equal], Project administration [equal], Writing – original draft [lead], Writing – review & editing [equal])

Supplementary data

Supplementary data is available at NAR online.

Conflict of interest

None declared.

Funding

This work was supported by the National Key R&D Program of China (Grant No. 2023YFA1800400 to G. Xu), the National Natural Science Foundation of China (32370625 and 32000420 to D.Z.), the Guangdong Major Project of Basic and Applied Basic Research (2023B0303000005 to G. Xu and P. Zhu). Funding to pay the Open Access publication charges for this article was provided by National Natural Science Foundation of China.

Data availability

All original data are available upon request.

References

- Cappannini A, Ray A, Purta E *et al.* MODOMICS: a database of RNA modifications and related information. 2023 update. *Nucleic Acids Res* 2024;52:D239–44. <https://doi.org/10.1093/nar/gkad1083>
- Raiber E-A, Hardisty R, van Delft P *et al.* Mapping and elucidating the function of modified bases in DNA. *Nat Rev Chem* 2017;1:0069. <https://doi.org/10.1038/s41570-017-0069>
- Pfaffeneder T, Spada F, Wagner M *et al.* Tet oxidizes thymine to 5-hydroxymethyluracil in mouse embryonic stem cell DNA. *Nat Chem Biol* 2014;10:574–81. <https://doi.org/10.1038/nchembio.1532>
- Jin GB, Rong SQ, Yin DR *et al.* Neomorphic leukemia-derived mutations in the TET2 enzyme induce genome instability via a substrate shift from 5-methylcytosine to thymine. *Proc Natl Acad Sci USA* 2025;122:e2418318122. <https://doi.org/10.1073/pnas.2418318122>
- Wu F, Li X, Looso M *et al.* Spurious transcription causing innate immune responses is prevented by 5-hydroxymethylcytosine. *Nat Genet* 2023;55:100–11. <https://doi.org/10.1038/s41588-022-01252-3>
- Parasyraki E, Mallick M, Hatch V *et al.* 5-Formylcytosine is an activating epigenetic mark for RNA Pol III during zygotic reprogramming. *Cell* 2024;187:6088–6103. <https://doi.org/10.1016/j.cell.2024.08.011>
- Kellinger MW, Song CX, Chong J *et al.* 5-Formylcytosine and 5-carboxylcytosine reduce the rate and substrate specificity of RNA polymerase II transcription. *Nat Struct Mol Biol* 2012;19:831–3. <https://doi.org/10.1038/nsmb.2346>
- Bilyard MK, Becker S, Balasubramanian S. Natural, modified DNA bases. *Curr Opin Chem Biol* 2020;57:1–7. <https://doi.org/10.1016/j.cbpa.2020.01.014>
- Xue JH, Chen GD, Hao F *et al.* A vitamin-C-derived DNA modification catalysed by an algal TET homologue. *Nature* 2019;569:581–5. <https://doi.org/10.1038/s41586-019-1160-0>
- Borst P, Sabatini R. Base J: discovery, biosynthesis, and possible functions. *Annu Rev Microbiol* 2008;62:235–51. <https://doi.org/10.1146/annurev.micro.62.081307.162750>
- Weigle P, Raleigh EA. Biosynthesis and function of modified bases in bacteria and their viruses. *Chem Rev* 2016;116:12655–87. <https://doi.org/10.1021/acs.chemrev.6b00114>
- Dou X, Boyd-Kirkup JD, McDermott J *et al.* The strand-biased mitochondrial DNA methylome and its regulation by DNMT3A. *Genome Res* 2019;29:1622–34. <https://doi.org/10.1101/gr.234021.117>
- Rebello AP, Williams SL, Moraes CT. *In vivo* methylation of mtDNA reveals the dynamics of protein-mtDNA interactions. *Nucleic Acids Res* 2009;37:6701–15. <https://doi.org/10.1093/nar/gkp727>
- Wu TP, Wang T, Seetin MG *et al.* DNA methylation on N(6)-adenine in mammalian embryonic stem cells. *Nature* 2016;532:329–33. <https://doi.org/10.1038/nature17640>
- Liu J, Zhu Y, Luo GZ *et al.* Abundant DNA 6mA methylation during early embryogenesis of zebrafish and pig. *Nat Commun* 2016;7:13052. <https://doi.org/10.1038/ncomms13052>
- Yao B, Cheng Y, Wang Z *et al.* DNA N6-methyladenine is dynamically regulated in the mouse brain following environmental stress. *Nat Commun* 2017;8:1122. <https://doi.org/10.1038/s41467-017-01195-y>
- Wang S, Xie H, Mao F *et al.* N(4)-acetyldeoxycytosine DNA modification marks euchromatin regions in *Arabidopsis thaliana*. *Genome Biol* 2022;23:5. <https://doi.org/10.1186/s13059-021-02578-7>
- Ye P, Luan Y, Chen K *et al.* MethSMRT: an integrative database for DNA N6-methyladenine and N4-methylcytosine generated by single-molecular real-time sequencing. *Nucleic Acids Res* 2016;45:D85–9. <https://doi.org/10.1093/nar/gkw950>
- Ratel D, Ravanat JL, Charles MP *et al.* Undetectable levels of N6-methyl adenine in mouse DNA: cloning and analysis of PRED28, a gene coding for a putative mammalian DNA adenine methyltransferase. *FEBS Lett* 2006;580:3179–84. <https://doi.org/10.1016/j.febslet.2006.04.074>
- Schiffers S, Ebert C, Rahimoff R *et al.* Quantitative LC-MS provides no evidence for m(6) dA or m(4) dC in the genome of mouse embryonic stem cells and tissues. *Angew Chem Int Ed* 2017;56:11268–71. <https://doi.org/10.1002/anie.201700424>
- Shao Z, Han Y, Zhou D. Optimized bisulfite sequencing analysis reveals the lack of 5-methylcytosine in mammalian mitochondrial DNA. *BMC Genomics* 2023;24:439. <https://doi.org/10.1186/s12864-023-09541-9>

22. Kong Y, Cao L, Deikus G *et al.* Critical assessment of DNA adenine methylation in eukaryotes using quantitative deconvolution. *Science* 2022;375:515–22. <https://doi.org/10.1126/science.abe7489>
23. Janulaitis A, Klimasauskas S, Petrusyte M *et al.* Cytosine modification in DNA by BcnI methylase yields N4-methylcytosine. *FEBS Lett* 1983;161:131–4. [https://doi.org/10.1016/0014-5793\(83\)80745-5](https://doi.org/10.1016/0014-5793(83)80745-5)
24. He YF, Li BZ, Li Z *et al.* Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science* 2011;333:1303–7. <https://doi.org/10.1126/science.1210944>
25. Kriaucionis S, Heintz N. The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science* 2009;324:929–30. <https://doi.org/10.1126/science.1169786>
26. Ohira T, Minowa K, Sugiyama K *et al.* Reversible RNA phosphorylation stabilizes tRNA for cellular thermotolerance. *Nature* 2022;605:372–9. <https://doi.org/10.1038/s41586-022-04677-2>
27. Wang Y, Nan B, Ye F *et al.* Dual modes of DNA N(6)-methyladenine maintenance by distinct methyltransferase complexes. *Proc Natl Acad Sci USA* 2025;122:e2413037121. <https://doi.org/10.1073/pnas.2413037121>
28. Kanaly RA, Hanaoka T, Sugimura H *et al.* Development of the adductome approach to detect DNA damage in humans. *Antioxid Redox Signal* 2006;8:993–1001. <https://doi.org/10.1089/ars.2006.8.993>
29. Guilbaud A, Ghanegolmohammadi F, Wang Y *et al.* Discovery adductomics provides a comprehensive portrait of tissue-, age- and sex-specific DNA modifications in rodents and humans. *Nucleic Acids Res* 2023;51:10829–45. <https://doi.org/10.1093/nar/gkad822>
30. Traube FR, Schiffers S, Iwan K *et al.* Isotope-dilution mass spectrometry for exact quantification of noncanonical DNA nucleosides. *Nat Protoc* 2019;14:283–312. <https://doi.org/10.1038/s41596-018-0094-6>
31. Shishodia S, Zhang D, El-Sagheer AH *et al.* NMR analyses on N-hydroxymethylated nucleobases – implications for formaldehyde toxicity and nucleic acid demethylases. *Org Biomol Chem* 2018;16:4021–32. <https://doi.org/10.1039/C8OB00734A>
32. Ammann G, Berg M, Dalwigk JF *et al.* Pitfalls in RNA modification quantification using nucleoside mass spectrometry. *Acc Chem Res* 2023;56:3121–31. <https://doi.org/10.1021/acs.accounts.3c00402>
33. Mo J, Liang Z, Lu M *et al.* Protonation-suppression-free LC-MS/MS analysis for profiling of DNA cytosine modifications in adult mice. *Anal Chem* 2020;92:7430–6. <https://doi.org/10.1021/acs.analchem.0c00962>
34. La Barbera G, Spenser Shuler M, Hammershøj Beck S *et al.* Development of an untargeted DNA adductomics method by ultra-high performance liquid chromatography coupled to high-resolution mass spectrometry. *Talanta* 2025;282:126985. <https://doi.org/10.1016/j.talanta.2024.126985>
35. Yin R, Mo J, Lu M *et al.* Detection of human urinary 5-hydroxymethylcytosine by stable isotope dilution HPLC-MS/MS analysis. *Anal Chem* 2015;87:1846–52. <https://doi.org/10.1021/ac5038895>
36. Guo C, Xie C, Chen Q *et al.* A novel malic acid-enhanced method for the analysis of 5-methyl-2'-deoxycytidine, 5-hydroxymethyl-2'-deoxycytidine, 5-methylcytidine and 5-hydroxymethylcytidine in human urine using hydrophilic interaction liquid chromatography-tandem mass spectrometry. *Anal Chim Acta* 2018;1034:110–8. <https://doi.org/10.1016/j.aca.2018.06.081>
37. Guo C, Hu Y, Cao X *et al.* HILIC-MS/MS for the determination of methylated adenine nucleosides in human urine. *Anal Chem* 2021;93:17060–8. <https://doi.org/10.1021/acs.analchem.1c03829>
38. Singh R, Teichert F, Seidel A *et al.* Development of a targeted adductomic method for the determination of polycyclic aromatic hydrocarbon DNA adducts using online column-switching liquid chromatography/tandem mass spectrometry. *Rapid Comm Mass Spectrometry* 2010;24:2329–40. <https://doi.org/10.1002/rcm.4645>
39. Sangaraju D, Villalta PW, Wickramaratne S *et al.* NanoLC/ESI+ HRMS3 quantitation of DNA adducts induced by 1,3-butadiene. *J Am Soc Mass Spectrom* 2014;25:1124–35. <https://doi.org/10.1007/s13361-014-0916-x>
40. Ma B, Jing M, Villalta PW *et al.* Simultaneous determination of 8-oxo-2'-deoxyguanosine and 8-oxo-2'-deoxyadenosine in human retinal DNA by liquid chromatography nanoelectrospray-tandem mass spectrometry. *Sci Rep* 2016;6:22375. <https://doi.org/10.1038/srep22375>
41. Iijima H, Patrzyc HB, Dawidzik JB *et al.* Measurement of DNA adducts in cells exposed to cisplatin. *Anal Biochem* 2004;333:65–71. <https://doi.org/10.1016/j.ab.2004.06.001>
42. Afoullouss S, Balsam A, Allcock AL *et al.* Optimization of LC-MS(2) data acquisition parameters for molecular networking applied to marine natural products. *Metabolites* 2022;12:245. <https://doi.org/10.3390/metabo12030245>
43. La Barbera G, Nommesen KD, Cuparencu C *et al.* A comprehensive database for DNA adductomics. *Front Chem* 2022;10:908572. <https://doi.org/10.3389/fchem.2022.908572>
44. Sood AJ, Viner C, Hoffman MM. DNAmod: the DNA modification database. *J Cheminform* 2019;11:30. <https://doi.org/10.1186/s13321-019-0349-4>
45. Xie Y, De Luna Vitorino FN, Chen Y *et al.* SWAMNA: a comprehensive platform for analysis of nucleic acid modifications. *Chem Commun* 2023;59:12499–502. <https://doi.org/10.1039/D3CC04402E>
46. Chang YJ, Cooke MS, Chen YR *et al.* Is high resolution a strict requirement for mass spectrometry-based cellular DNA adductomics? *Chemosphere* 2021;274:129991. <https://doi.org/10.1016/j.chemosphere.2021.129991>
47. Hu C-W, Chang Y-J, Chang W-H *et al.* A novel adductomics workflow incorporating FeatureHunter software: rapid detection of nucleic acid modifications for studying the exposome. *Environ Sci Technol* 2024;58:75–89. <https://doi.org/10.1021/acs.est.3c04674>
48. Kellner S, Neumann J, Rosenkranz D *et al.* Profiling of RNA modifications by multiplexed stable isotope labelling. *Chem Commun* 2014;50:3516–8. <https://doi.org/10.1039/c3cc49114e>
49. Kellner S, Ochel A, Thuring K *et al.* Absolute and relative quantification of RNA modifications via biosynthetic isotopomers. *Nucleic Acids Res* 2014;42:e142. <https://doi.org/10.1093/nar/gku733>
50. Tozzi MG, Camici M, Mascia L *et al.* Pentose phosphates in nucleoside interconversion and catabolism. *FEBS J* 2006;273:1089–101. <https://doi.org/10.1111/j.1742-4658.2006.05155.x>
51. Mushev MU, Baumgartner A, Krebs L *et al.* The origin of genomic N(6)-methyl-deoxyadenosine in mammalian cells. *Nat Chem Biol* 2020;16:630–4. <https://doi.org/10.1038/s41589-020-0504-2>
52. Gupta R, Gupta N. *Fundamentals of Bacterial Physiology and Metabolism*. Singapore: Springer Singapore, 2021; pp.525–54. <https://doi.org/10.1007/978-981-16-0723-3>
53. Liu X, Lai W, Li Y *et al.* Author correction: n(6)-methyladenine is incorporated into mammalian genome by DNA polymerase. *Cell Res* 2021;31:241. <https://doi.org/10.1038/s41422-020-0339-0>
54. Lyu C, Niu Y, Lai W *et al.* Rare and misincorporated DNA N(6)-methyladenine is a hallmark of cytotoxic stresses for selectively stimulating the stemness and proliferation of glioblastoma cells. *Cell Discov* 2022;8:39. <https://doi.org/10.1038/s41421-022-00399-x>
55. Stincone A, Prigione A, Cramer T *et al.* The return of metabolism: biochemistry and physiology of the pentose phosphate pathway. *Biol Rev* 2015;90:927–63. <https://doi.org/10.1111/brv.12140>

56. O'Brien ZK, Boulias K, Wang J *et al.* Sources of artifact in measurements of 6mA and 4mC abundance in eukaryotic genomic DNA. *BMC Genomics* 2019;20:445.
<https://doi.org/10.1186/s12864-019-5754-6>
57. Liu B, Liu X, Lai W *et al.* Metabolically generated stable isotope-labeled deoxynucleoside code for tracing DNA N(6)-methyladenine in human cells. *Anal Chem* 2017;89:6202–9.
<https://doi.org/10.1021/acs.analchem.7b01152>
58. De Graeve M, Van de Walle E, Van Hecke T *et al.* Exploration and optimization of extraction, analysis and data normalization strategies for mass spectrometry-based DNA adductome mapping and modeling. *Anal Chim Acta* 2023;1274:341578.
<https://doi.org/10.1016/j.aca.2023.341578>
59. Beland FA, Fullerton NF, Heflich RH. Rapid isolation, hydrolysis and chromatography of formaldehyde-modified DNA. *Journal of Chromatography B: Biomedical Sciences and Applications* 1984;308:121–31.
[https://doi.org/10.1016/0378-4347\(84\)80202-9](https://doi.org/10.1016/0378-4347(84)80202-9)
60. Zhong W, Hee SS. Comparison of UV, fluorescence, and electrochemical detectors for the analysis of formaldehyde-induced DNA adducts. *J Anal Toxicol* 2005;29:182–7.
<https://doi.org/10.1093/jat/29.3.182>
61. Ma J, Qi R, Harcourt EM *et al.* 3,N4-Etheno-5-methylcytosine blocks TET1-3 oxidation but is repaired by ALKBH2, 3 and FTO. *Nucleic Acids Res* 2024;52:12378–89.
<https://doi.org/10.1093/nar/gkac818>
62. Pham VN, Bruemmer KJ, Toh JDW *et al.* Formaldehyde regulates S-adenosylmethionine biosynthesis and one-carbon metabolism. *Science* 2023;382:eabp9201.
<https://doi.org/10.1126/science.abp9201>
63. Kanaly RA, Hanaoka T, Sugimura H *et al.* Development of the adductome approach to detect DNA damage in humans. *Antioxid Redox Signal* 2006;8:993–1001.
<https://doi.org/10.1089/ars.2006.8.993>
64. Guilbaud A, Ghanegolmohammadi F, Wang Y *et al.* Discovery adductomics provides a comprehensive portrait of tissue-, age- and sex-specific DNA modifications in rodents and humans. *Nucleic Acids Res* 2023;51:10829–45.
<https://doi.org/10.1093/nar/gkad822>
65. Walters K, Stornetta A, Jacobs F *et al.* Identification of new candidate biomarkers to support doxorubicin treatments in canine cancer patients. *BMC Vet Res* 2021;17:378.
<https://doi.org/10.1186/s12917-021-03062-x>